# Music Recommendation System Based on User's Sentiments Extracted from Social Networks

Renata L. Rosa, Demóstenes Z. Rodríguez, and Graça Bressan

**Abstract** —*In recent years, the sentiment analysis has been explored by several Internet services to recommend contents in accordance with human emotions, which are expressed through informal texts posted on social networks. However, the metrics used in the sentiment analysis only classify a sentence with positive, neutral or negative intensity, and do not detect sentiment variations in accordance with the user's profile. In this arena, this paper presents a music recommendation system based on a sentiment intensity metric, named enhanced Sentiment Metric (eSM) that is the association of a lexicon-based sentiment metric with a correction factor based on the user's profile. This correction factor is discovered by means of subjective tests, conducted in a laboratory environment. Based on the experimental results, the correction factor is formulated and used to adjust the final sentiment intensity. The users' sentiments are extracted from sentences posted on social networks and the music recommendation system is performed through a framework of low complexity for mobile devices, which suggests songs based on the current user's sentiment intensity. Also, the framework was built considering ergonomic criteria of usability. The performance of the proposed framework is evaluated with remote users using the crowdsourcing method, reaching a rating of 91% of user satisfaction, outperforming a randomly assigned song suggestion that reached 65% of user satisfaction. Furthermore, the paper presents low perceived impacts on the analysis of energy consumption, network and latency in accordance with the processing and memory perception of the recommendation system, showing advantages for the consumer electronic world.* [1]

**Index Terms — Recommendation System, Social Network, Mobile Devices, Sentiment Analysis.**

## I. Introduction

People use the Internet to express themselves and social networks have become a popular way to share information, ideas and experiences. People use social networks to write sentences with positive, negative or neutral emotions, expressing their feelings; in this context, studies concerning sentiment intensity have started to emerge. Knowledge of the sentiment intensity of a sentence can help to collect useful information and allows to know more about a person who expresses herself or himself about an event, product or content.

Sentiment analysis is a technique of natural language processing and text analytics, which can be applied to many areas, such as e-learning, e-commerce [1], and multimedia [2] among others. However, its use in recommendation systems remains a challenge; people express their feelings in different ways, making it difficult to create reliable recommendations based on sentiments [3].

Sentiment analysis is starting to be explored in music recommendation systems to suggest a specific song depending on the emotional state of a person, since the song is totally related to the current emotion and feelings of the person. There is sentiment analysis research based on physiological signals [4], [5], subjective emotion assessment [6], tag-based extractions [7], [8], web semantic [9], [10], machine learning, such as, Support Vector Machines (SVM) and its derivation [11], and the lexicon-based technique such as the ANEW [12].

A user of a music recommendation system can choose his or her emotional state manually [6] using emoticon faces or by physiological sensors in a music recommendation system [5]. However, this collection of emotions makes the system dependent on the user input and equipment. It is therefore ideal to have an independent and automatic recommendation system.

Tag-based recommendations and lexicon-based techniques [13] aim to extract emotions in accordance with words, not differentiating whether or not a person is more emotional. Consequently, there is a gap in current studies in how people express themselves with greater or lesser intensity of feeling.

People use social networks at certain times of day [14] and knowing their habits and behaviors may be useful for several applications; therefore, recording these data in logs is very important. By the implementation of logs, it is possible to develop a recommendation for a customized period of time. For instance, a person usually reads financial news, e.g., from 9a.m. to 10a.m. on Monday and Tuesday and tends to read entertainment news on Saturday [14].

A recommendation system can collect data all the time and disseminate a large amount of information to users, but that is an inconvenience for the system and its users, in addition to consuming more processing, bandwidth, and memory resources of a consumer electronic device. Therefore, widely used applications must consume fewer resources in times of an increased number of consumers' portable devices.

This paper is an extended contribution to the work [2], which proposes a personal music recommendation system

[1] R. L. Rosa, and G. Bressan are with the Computer Engineering Department at the School of Engineering - University of São Paulo, CEP 5508-010, Brazil (e-mails: rrosa@usp.br and gbresan@larc.usp.br).

D. Z. Rodríguez is with the Computer Science Department at the Federal University of Lavras, CEP 37200-000, Brazil (e-mail: demostenes.zegarra@dcc.ufla.br).

Contributed Paper
Manuscript received 06/30/15
Current version published 09/29/15
Electronic version published 09/29/15.

0098 3063/15/$20.00 © 2015 IEEE

using a new lexicon-based sentiment metric named *eSM*, which uses the Sentimeter-Br2 [2] metric associated with a novel correction factor based on the user's profile, to be applied to the sentiment analysis.

The Sentimeter-Br2 metric is a word dictionary with respective sentiment intensity of positive or negative value, which considers n-grams, adverbs, removes stopwords, words which do not add sentiment to a sentence. Also, the metric differentiates sentiment values depending on verbal tenses, in which a verb in the past tense has a lesser sentiment value than a verb in the present tense. Sentimeter-Br2 is based on the Sentimeter-Br [1].

The main goal of this paper is to demonstrate that a lexicon-based sentiment intensity metric associated to a correction factor can improve the performance of a music recommendation system, using a low complexity solution with advantages to the consumer electronic devices. The correction factor uses the characteristics of a person that can easily be extracted from social networks. The new metric, *eSM*, therefore presents a more accurate sentiment value.

Additionally, the proposed recommendation system considered ergonomic factors of usability in order to improve the user's quality of experience.

In this context, the remainder of this paper is structured as follows. Section II presents the related work. Section III presents the proposed sentiment intensity model used to determine the *eSM*. Section IV shows the proposed recommendation system architecture based on sentiments. Section V presents the experimental results that include the proposed recommendation system with the *eSM* performance evaluation. Finally, section VI presents the conclusions.

## II. RELATED WORK

In this section, firstly, the main studies regarding sentiment analysis are reviewed. Secondly, the recommendation systems using metrics of sentiment analysis are treated.

### *A. Sentiment Analysis*

Sentiment analysis can be performed using three approaches, as shown in Fig. 1, the corpus-based approach using machine learning [15], the lexicon-based approach using a word dictionary [16] and a hybrid-based method which combines both approaches [17].

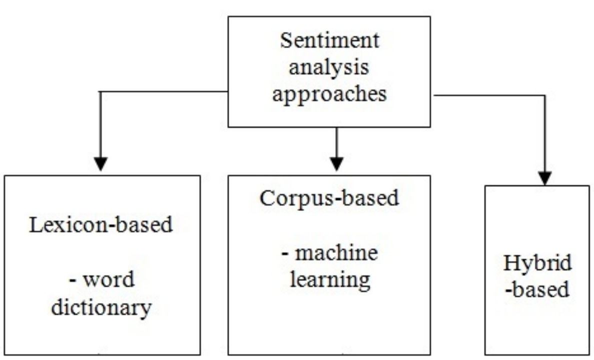


**Fig. 1. Different approaches for performing the sentiments analysis.**

These approaches are used to perform sentiment analysis. However, notice that the use of machine learning requires a large amount of data to obtain a reliable sentiment result, because an unusual sentence may cause noise in the calculation of the sentiment.

This research focuses on the lexicon-based approach using a word dictionary, which is used to define the sentiment intensity metric. Many studies of sentiment analysis try to improve the metric to find a more accurate sentiment intensity of a text. A manual dictionary consists of words, in which each word has a respective classification, e.g., a positive scale from +1 to +5 or a negative scale from -1 to -5 [12], as shown in Table I. A dictionary of affective language [18] uses an intensity scale with many emotional words. WordNet is another dictionary of words that can be used for sentiment analysis. In this study, both the SentiStrength [19] and the Sentimeter-Br2 metrics are used, because they both support the Portuguese language and a comparison of their performance can be made.

TABLE I
WORDS FROM A DICTIONARY WITH RESPECTIVE SENTIMENT INTENSITY VALUES

| Word | Sentiment intensity value |
|---|---|
| bad | -2 |
| like | +2 |
| combat | -3 |
| beautiful | +3 |

Once the dictionary to be used is defined, the sentiment intensity metric can be modeled. The basic metric to obtain the sentiment of a sentence is commonly obtained by an arithmetic sum of each word found in the dictionary. For example, in (1), $Sentiment\ (F)$ represents the total sentiment of sentence F, and the variable $dictionary\ .value$ is the sentiment value defined in the dictionary of each word $W_i$ that composing the sentence.

$$Sentiment(F) \ = \sum_{i=1}^{m} dictionary.value(W_i) \quad (1)$$

In the sentence "I like beautiful shoes" only the words "like" and "beautiful" have sentiment intensity values, of +2 and +3, respectively, in a word-based dictionary. Using (1), the result of the total sentiment of this sentence in (2) is equal to +5; therefore, the sentence has a positive sentiment intensity value:

$$Sentiment(F) \ = \sum_{i=1}^{m} \ +2+(+3)=+5 \quad (2)$$

In order to calculate the sentiment intensity of more complex sentence it is necessary to identify the verb tenses, adverbs and n-grams. This study uses a metric of sentiments based on [1], because this metric considers different words, grammatical classes and verb tenses.

Sentiment analysis is beginning to be explored in many fields, but there are many points to be studied, such as whether the user's profile really influences a sentiment metric, the characteristics of which must be considered and how to perform the association between the user's profile and the sentiment intensity metric. For example, Thelwall *et al.* [20] states that the sentiment intensity may vary depending on gender, but the association with a metric was not defined.

Following this general review of sentiment analysis metrics, it can be concluded that the lexicon-based metrics do not differentiate the sentiment depending on gender, age or other user's characteristic.

*B. Recommendation Systems*

There are three recommendation systems approaches, content-based, collaborative and hybrid-based. The content-based approach works with the association between the description of an item and the user's profile; the suggestion of items is based on the user's preferences. The collaborative-based approach analyses the user behavior and preferences and explores similar preferences between people [21], [22]. The hybrid approach combines both methods.

A recommendation system is allocated on a server [23] containing a database with users' preferences or history. The advantages of using the server allocation are that they do not overload the memory of the user's device with data, since the client-side application is of low complexity. There is a data transmission flow between the client and server, but the current data network is not as restricted as it used to be.

In the recommendation systems, the sentiment analysis began to be explored to suggest more updated contents and based on the person's mood and feelings.

Tag-based recommendation [7], [8], [13] focuses on music use suggestions to search for words and count the word frequency; the system recommends songs containing the words that were searched by people. However, the recommendations are only based on words already searched by the user, limiting the suggestions of new words.

The web semantic recommendations consider syntactic similarity metrics to offer similar preferences. However, more intelligent techniques are necessary to offer contents in a more flexible way, related to user's preference and semantic associations [10]. Content-based recommendations consider the relations between words and ontology to recommend contents and not the user's feeling variation for recommending contents.

Recommendation systems in specific areas such as in multimedia need to suggest contents based on the person's current mood, emotion and feeling because the person chooses a particular film or song depending on their current mood and feeling.

The user feedback in any recommendation system is an important tool to be considered. Studies which focus on digital TV or music recommendations make suggestions according to user feedback to be more effective [24]. Users do not often like to write manual feedback; it is therefore important to use an automatic technique to collect user feedback. Sentiment analysis is a way of collecting users' satisfaction feedback automatically.

The user's profile is another important factor that must be included in a recommendation system, as collected by a questionnaire or rating [1], [24]. On social networks it is possible to capture some user profile details, such as gender, age, educational level and city of birth, automatically. 93.8% of people fill the gender category [25] and almost 70% of people fill in their country, education and age on social networks [26].

Studies on music recommendation systems based on emotions [6] usually suggest songs depending on a manual evaluation of emotion through emoticon faces, as shown in Fig. 2. However, for a system to represent the person's mood swings over the course of a day, one would have to choose their mood manually and daily in the morning, afternoon and evening. [20].

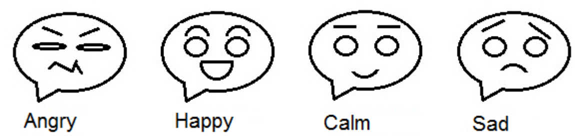


**Fig. 2. Example of emoticon faces used on music recommendation systems.**

It is possible to obtain the user's emotion automatically through sensors [4], [5]. The equipment for measuring emotions includes electroencephalogram signals and electrodes for the acquiring peripheral physiological signals [5]. However, people do not generally have a sensor to measure physical parameters to constantly transmit this information by Internet.

Ideally, current recommendation systems calculate the sentiment intensity of sentences extracted from real social networks as opposed to extracting sentences from a specific trusted network [27].

As stated before, most of the recommendation systems send a suggestion based on both variables, based on the user history and on the user's profile, but nowadays, the recommendations are starting to explore making more suggestions based on sentiment analysis. The possibility of recognizing if a person is happy or sad and suggesting content in accordance to their current mood state help to discover when the person is more likely to receive suggestions; allowing the user to feel less flooded with general or past suggestions.

## III. The Proposed Sentiment Intensity Metric

The proposed sentiment model is presented in this section. The model shows that if a sentiment metric does not consider the user's profile, the metric is not able to generate a real sentiment intensity value.

A correction factor, which is discovered by subjective tests, needs to be applied in accordance with the user's profile on social networks. The subjective tests were performed in two phases; firstly, in a laboratory environment to find which parameters of the user's profile affect the sentiment value of a sentence and to evaluate the performance and cognitive factors of the framework. In the second phase, a remote subjective method was used to validate the proposed model. The tests conducted in a laboratory served to test initial theories, and the remote method served to validate these theories.

### *A. Initial Studies using Subjective Tests in a Laboratory Environment*

At that stage, initial subjective tests were performed in a laboratory environment by Portuguese native-speaking assessors. In the tests, it was possible to work with a wider diversity of assessor profiles, such as region of birth (North, Northeast, Midwest, Southeast and South of Brazil), religion, race, educational level, among other characteristics discovered through the web interface questionnaire presented in Fig. 3. Furthermore, a question about which musical genre the assessor prefers was asked in the questionnaire, based on his/her emotional state (sad, happy and calm); the person could choose one or two musical genre options.

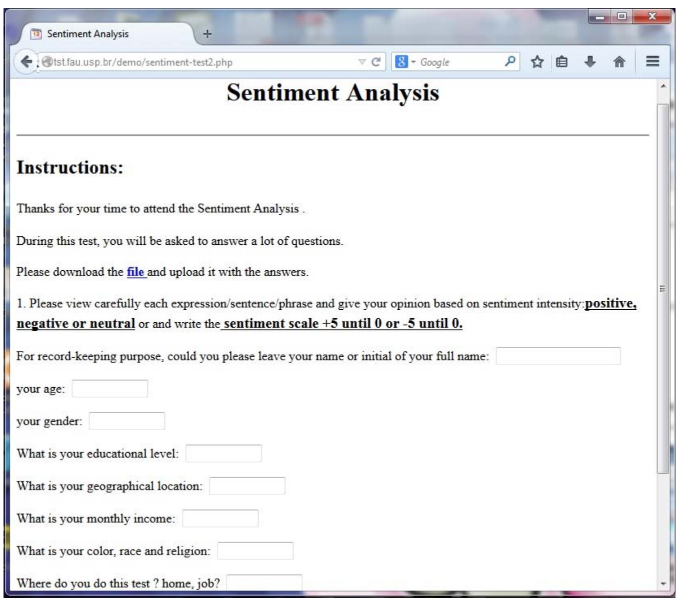


**Fig. 3. Web interface available for the assessors to answer the questions.**

The initial characteristics extracted from Fig. 3 helped to know what person's characteristics could affect a sentiment metric.

The assessors' answers were collected and the field corresponding to age was separated into four groups.

The studies were performed by 200 assessors, consisting of 100 women and 100 men; each participant wrote his or her musical preference and filled out a questionnaire with his or her profile. Later, each person posted sentences on social networks, which were captured by a script. These sentences were then analyzed by both the same person who posted the sentences and by the Sentimeter-Br2 metric. The person evaluated each sentence on a sentiment scale from +5 to -5 in the first phase of the tests.

The assessors were monitored to ensure they collected all the sentences they posted on social networks every hour over a test period of 3 weeks. After the first day, the tests were carried out from a distance because people were not available to be at the laboratory for several hours at a time during the test period. The social network username was already known and the script captured the user's sentences automatically. After 3 weeks, all the captured sentences were analyzed by both the Sentimeter-Br2 and the assessor who posted the sentences.

At the end of the 3-week period, the assessors went to the laboratory again to test the performance of the framework. Over a 15-day period, each assessor chose a day to evaluate the application through tests with a standard mobile phone.

In total, 19,600 sentences were extracted from the social network and evaluated using the scale shown in Table II. Out of the 19,600, only 652 sentences were discarded because they were considered spam.

Access logs were collected to study the users' routines. In the experiments, the observation that each user had a custom period of time to access and post sentences on social networks was made, with a window of 5 to 20 minutes; for instance, user A preferred to post sentences between at 10:00 pm with a window of 15 minutes, and later, user A usually posted more sentences between 9:45 pm to 10:15 pm.

The average value of the access time to the social network is added to the music recommendation framework in order to capture the sentences of each user. This information is useful for capturing sentences in a customized period of time instead of randomly and constantly capturing the sentences. The framework thereby saves more processing memory and energy resources.

A sentiment correction factor based on the user profile was obtained, through the subjective tests. The mathematical model represents a correction factor to be applied to traditional sentiment metrics.

### *B. The Proposed Correction Factor obtained through User's Profile*

The results of the initial studies using subjective tests helped to create the proposed mathematical model of the correction factor (CF) that adjusts Sentimetr-Br2 giving the new *eSM* metric as a result, as shown in Fig. 4.

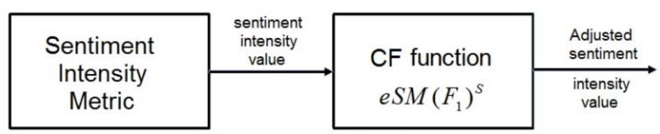


**Fig. 4. Sentiment correction factor (CF) function used to adjust the sentiment intensity obtained by a sentiment intensity metric.**

The $eSM(F_1)^S$ value of a sentence $F_1$ is given by (3). The $eSM(F_1)^S$ represents the sentiment intensity value of the sentence $F_1$. This value was obtained from the subjective test results, which represent the sentiment intensity value given by the assessors. The $eSM(F_1)$ is the sentiment intensity value of the sentence determined by the Sentimeter-Br2 metric.

$$eSM(F_1)^S = SM(F_1) * C * \exp(a_1 * A_1 + a_2 * A_2 + \ldots + a_n * A_n + g_1 * M + g_2 * F + \mathrm{e1} * \mathrm{G} + \mathrm{e2} * \mathrm{nG}) \quad (3)$$

Where: C is a scale constant; $a_1 \ldots a_n$ are binary factors related to age ranges, if one of them is equal to one, the others are zeros; $A_1 \ldots A_n$ are the weight factors of each age range. This paper considers four ranges; $g_1$ and $g_2$ are binary factors related to the gender, if one of them is equal to one, the other is zero; M and F are the weight factors of gender, man or woman, respectively; $e_1$ e $e_2$ are binary factors related to educational level (higher education or not); G e nG are the weight factors of educational level, higher education or not, respectively.

It is worth noting that all the parameters presented in Table II were not considered in (3), because subjective test results demonstrated that they do not affect the $eSM(F_1)^S$ value.

The linear and exponential functions were tested, and the last function presented the lowest squared error.

The metric Sentimeter-Br2 was considered in this paper, but any sentiment intensity metric can be used to calculate $eSM(F_1)$ in (3).

## IV. Proposed Recommendation System based on Sentiments

This section covers the details and methodology of the proposed music recommendation system based on the *eSM* sentiment metric - the system architecture is introduced in Fig. 5.

The mathematical model obtained in (3) is applied to the proposed recommendation system to recommend music based on the user's sentiments in a customized period of time.

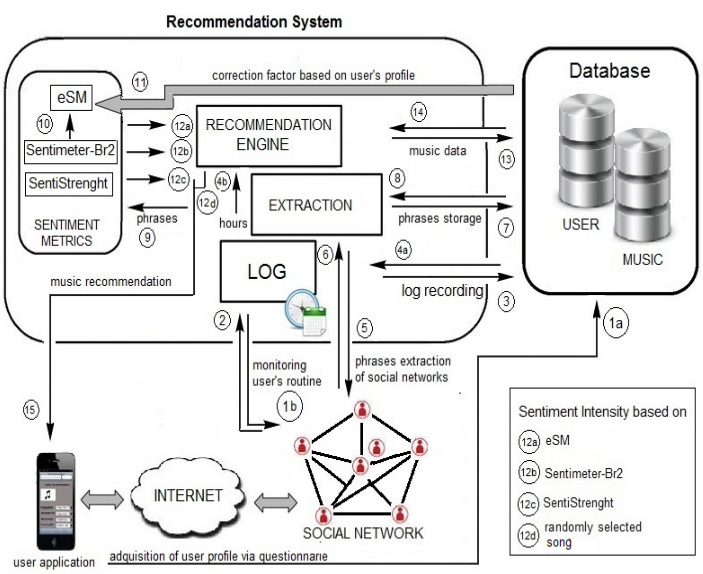


**Fig. 5. Framework architecture of the proposed recommendation system based on sentiment intensity metrics.**

The recommendation system has a user's profile database, which contains the user's musical preference, the profile parameters as age, gender, educational level, musical preference based on his/her emotional state and the social network username of the person who completed the questionnaire on the system for the first time. Initially, the system captures every sentence posted by a person; after three weeks the system has logged the times of day that the person commonly uses the social network. The system thereby only captures sentences at that time of day or week, with a time window of ±20 min.

As can be observed in Fig. 5, the framework follows four recommendation system models. Firstly, the acquisition of user's profile is performed and the data is sent to the database (1a); the monitoring of user's routine starts(1b) and (2), in which the time and date is recorded in a log and sent to the database (3) and to the recommendation engine (4a) and (4b); the user's sentence are extracted from the social networks (5) and the sentences are sent to the extraction module engine (6); later, the sentences are stored in the database (7) and sent to the metrics to calculates the sentiment intensity (8) and (9). The sentiment intensity values are calculated by the metrics and the sentiment associated to the user's profile (11).The *eSM* is calculated using (10) and (11). The 3-sentiment metrics (12a), (12b), (12c) and the randomly selected song (12d) are sent to the Recommendation Engine, which requires (13) the selected songs and receives them (14). Finally, these song recommendations are sent to the user application (15).

If no sentence is posted on social networks, then songs of his or her preference style are recommended.

### A. Databases

240 songs are extracted from a Brazilian music portal and categorized in sentiment intensity and emotions by specialists in music; the songs are stored as file streams in MPEG-I Layer III audio coding scheme known as MP3 in the database to be used in the recommendation system.

The musical categories used in the recommendation system are based on happy, sad, angry and calm emotions [28], shown in Fig. 6, which are based on Russell's circumplex model [29] of affect with arousal and valence dimensions. Each song used in the framework was classified in one or two categories, for example, a song can be considered calm and sad, at the same time.

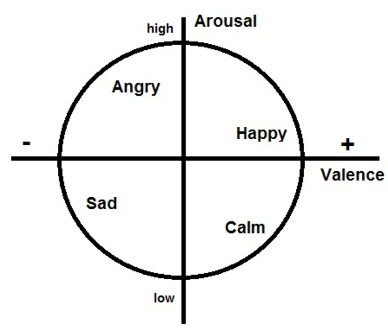


**Fig. 6. Plot of emotions used to categorize the musical genres: happy, sad, angry and calm.**

A set of records containing the name of the song, style, singer, sentiment intensity and emotion is stored in the database. The majority of the songs have 3.0 MB as maximum size and the average length is in the range of 2 to 3 minutes. The following types of music were considered: sad, happy and calm.

The database also stores the user's profile, social network username, category and the user's favorite song. The data is stored in an Open Source database.

### B. Client Application

The programming languages used in the client and in the server-side applications are Open Source languages. The client web application is written in the Hypertext Preprocessor (PHP), JavaScript Object Notation (JSON) and HTML5.

The web interface made in PHP and HTML5 is presented to the person with the song suggestion; the music emotion classification, song name, singer and the song is presented on the interface. As shown in Fig. 7, the person reads the music suggestion name according to the four recommendations and clicks on the plus symbol to listen the recommended song.


**Fig. 7. Main menu screen presented in the mobile device with four recommendation models.**

### C. Server

The framework has the modules engine of recommendation, sentence extraction and database storage. The server uses the Apache Web Server and PHP modules. The PHP language is a server-side programming language and communicates with a relational database to store the data about the user and music.

The sentences are extracted from the social network by an automatic script written in the PHP language and JSON. The databases and the engine of the sentiment intensity metric occur on the server side, so that the user's device is not overloaded, and only needs a network communication to transfer data.

Each time the user writes sentences on social networks, the server collects the sentences in the period of time registered in the log and sends the suggestions in another period of time. For instance, user A only posts 4 sentences a day and user B prefers to post sentences every 10 minutes throughout the day. The songs are suggested every 20 minutes for both users, but with a limit of 12 suggestions in the morning, afternoon and evening. This is because the person's emotion stabilizes for a few hours of the day, normally changing from one period of the day to another, for example, changing from morning to the afternoon [19].

### D. Evaluation Tests using Remote Assessors

Tests conducted by remote assessors are valid to evaluate the performance of framework and applications [30]-[32]. Remote tests can be performed though commercial platforms [33], in which workers perform a task and receive a monetary compensation, or solutions that use voluntary workers.

For this project, the assessors were selected from a commercial *crowdsourcing* portal. 300 remote Portuguese-speaking users were selected to answer the questions of Table II, which parameters are used in $eSM(F_1)$ and in the recommendation system. Also, they evaluated the performance of the proposed recommendation system.

TABLE II
PARAMETERS COLLECTED FROM THE ASSESSORS

| Field | Kind |
|---|---|
| Gender | man, woman |
| Educational level | higher education or not |
| Rating scale to score the posted sentences | +0.1 until +5 (positive sentences),<br>-0.1 until -5 (negative sentences)<br>or neutral sentences |
| Musical preference based on his/her emotional state | sad emotional state: sad, happy or calm<br>happy emotional state: sad, happy or calm<br>calm emotional state: sad, happy or calm |

Later, the assessors answered the questionnaire; they were separated into different groups taking into account age, gender and educational level.

## V. Experimental Results

Subjective tests were performed in a laboratory environment proposing to determine the $eSM(F_1)^S$ metric, the perceived resources consumed in the electronic device and the usability of the mobile application. Finally, the remote method was used to validate the performance of the $eSM(F_1)^S$, in which the satisfaction in accordance with the recommendation messages is measured. The results are explained as follows.

### A. Subjective Tests in a Laboratory Environment

Assessors went to a laboratory and the tests were conducted individually with no disturbing noises. At the beginning of the tests, an explanation was given stating that each assessor should fill in a questionnaire with his or her user's profile, post sentences and score each sentence. At the end of 3 weeks, the assessor should evaluate the framework performance.

Fig. 8 presents the weight of each parameter considered in (3). The results show that the gender and age parameters are the most influential factors in a sentiment intensity metric. Notice that the recommendation system needs to extract simple user's characteristics, as age, gender, educational level and song preferences based on the user's emotion. The application does not need to extract complex user's characteristics, as monthly income, religion and race because these characteristics do not affect the sentiment analysis. Therefore, the study does not violate any privacy rule and can be applied in any country.

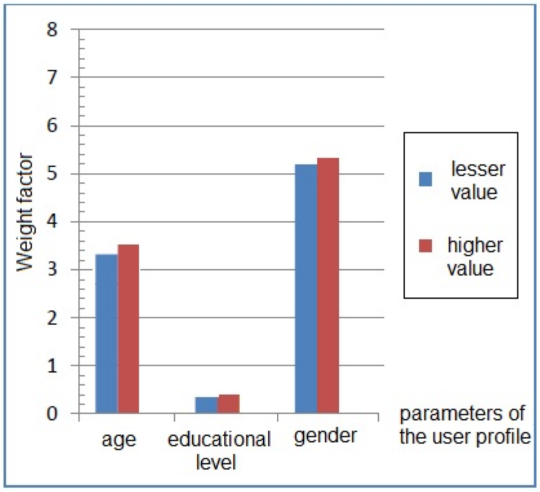


**Fig. 8. Weight factors obtained by subjective tests performed in laboratory on the parameters: gender, age and educational level.**

Assessors evaluated the perceived consumed resources from the recommendation system application running in the electronic device and presented in Table III. The scores are based on a 5-point scale, in which a grading value of 5 means excellent quality and an imperceptible interruption and a grading value of 1 means bad quality or a very annoying interruption.

The high average values of parameters presented in Table III show that the person perceives that the application consumes low resources in the electronic device.

TABLE III
PERCEIVED VALUE OF CONSUMED RESOURCES IN THE ELECTRONIC DEVICE AND RESPECTIVE EVALUATION IN ACCORDANCE WITH A 5-POINT SCALE

| Performance Regarding the Parameters | Perceived Average Value |
|---|---|
| Latency of the recommendation system in general | 4.2 |
| Energy consumption of the recommendation system application | 4.2 |
| Apparent network resource consumption | 4.5 |

The recommendation system application was tested using a mobile device with a Wireless Fidelity interface (Wi-Fi) to transfer the data between the client and the server. The device characteristics are shown in Table IV.

The application was built to be performed in 1 core processor of 1.6 GHz octa-core because it is a simple code to be applied in a recommendation system.

Some usability parameters [34] that consider extrinsic and intrinsic properties [35] were also analyzed by the assessors in relation to the framework. The 5-point scale average value of the parameters is shown in Table V. These usability properties express the transparency of the operative interface [36] of the framework.

TABLE IV
SPECIFICATIONS OF THE ELECTRONIC DEVICE USED IN THE TESTS

| Specifications | Values |
|---|---|
| Battery Capacity | 2,600 mAh |
| Touchscreen | Yes |
| Display Resolution | 1080 x 1920 pixels |
| Display Size | 5.0-inch touchscreen - Full HD |
| Colours | 16M colors |
| Processor | 1.6 GHz octa-core |
| Random Access Memory (RAM) | 2Gb |
| Internal data storage | 16Gb |

The usability parameters tested in the proposed framework included: (1) the prompting parameter that refers to all the means that help users to recognize the alternatives when several actions are presented in the system; (2) the immediate feedback that corresponds to a fast response from the system with the information provided in the requested transaction; (3) the legibility, in which the information presented on the screen must facilitate the reading of this information; (4) the concision that concerns perceptual workload for individual inputs or outputs; (5) the user control, the system's users should always be in control of the system; (6) the flexibility that reflects possible ways of achieving a given goal; (7) the user experience that refers to the means available to take into account the level of user experience in the system; (8) the error protection that detects and prevents errors in general; (9) the consistency that refers to the fact that information should always be presented in similar procedures to access the menu options.

TABLE V
PERCEIVED VALUE OF USABILITY PARAMETERS OF THE FRAMEWORK IN ACCORDANCE WITH A 5-POINT SCALE

| Usability Parameters | Perceived Value |
|---|---|
| Prompting | 4.9 |
| Immediate Feedback | 4.6 |
| Legibility | 4.8 |
| Concision | 4.2 |
| User Control | 4.1 |
| Flexibility | 4.4 |
| User Experience | 3.9 |
| Error Protection | 4.1 |
| Consistency | 4.7 |

The goal of using the usability parameters in an application assessment is to show that an application must consider parameters of ergonomic criteria of usability without consuming many resources from the electronic devices. The ergonomic parameters are therefore important to present a useful and easy system platform to people.

### *B. Performance Evaluation of the Proposed Recommendation System*

The performance evaluation of the proposed recommendation system was carried out by the remote assessors.

Table VII shows the results of the users' satisfaction level with the recommendation systems. The answer options are in accordance with a scale based on adjectives described in Likert [37], which are: very good, good, neutral, poor and very poor. This scale represents a qualitative measurement and has been widely applied to many studies [38], [39].

The users evaluated the recommendation system positively using the *eSM* metric. The results reached 91% users satisfaction with the framework using the proposed metric *eSM*.

**TABLE VII**
**% OF LIKERT ANSWERS OF THE RECOMMENDATION USING METRICS OF SENTIMENT AND RANDOMLY ASSIGNED SONG**

| | Randomly assigned song suggestion (no sentiment metric) | Sentimeter-Br2 | SentiStrengh | *eSM* (proposed metric) |
|---|---|---|---|---|
| Very good | 65% | 78% | 70% | 91% |
| Good | 15% | 13% | 16% | 7% |
| Neutral | 10% | 6% | 4% | 1% |
| Poor | 8% | 2% | 8% | 1% |
| Very poor | 2% | 1% | 2% | 0% |

The correction factor initially applied to Sentimeter-Br2 metric to obtain *eSM* as shown in (3), was also applied to the SentiStrengh metric. The results are presented in Table VIII. It is worth noting that another sentiment intensity metric can be used.

**TABLE VIII**
**% OF LIKERT ANSWERS OF THE RECOMMENDATION USING THE SENTIMENT *eSM* METRIC WITH THE SENTISTRENGH AND SENTIMETER-BR2**

| | *eSM* (using SentiStrengh metric) | *eSM* (using Sentimeter-Br2 metric) |
|---|---|---|
| Very Good | 84% | 91% |
| Good | 9% | 7% |
| Neutral | 5% | 1% |
| Poor | 2% | 1% |
| Very Poor | 0% | 0% |

## VI. CONCLUSION

The subjective tests results highlight the importance of considering the user's profile in a sentiment metric. Thus, the tests in the laboratory environment have shown what parameters may influence the final sentiment intensity of a sentence. Based on the results, a correction factor was obtained, which depend on age, educational level and gender. The correction factor was used to obtain a more real sentiment intensity value.

The new sentiment intensity metric, *eSM*, improved the music recommendation system, showing that the sentiments can change depending on the user's profile. The tests showed that the weight factor used in the *eSM* can be applied to another sentiment metric, for instance, the SentiStrengh metric.

The remote subjective tests reached 91% user satisfaction regarding the *eSM* in contrast to 65% of a randomly assigned song suggestion that did not consider a sentiment intensity, and 78% user satisfaction was reached by considering only a sentiment intensity metric, the Sentimeter-Br2.

The users profile were analyzed and the results showed that 78% of users preferred to listen to a musical genre similar to their current emotional state, and only 22% preferred to listen to a different musical genre in relation to their current emotional state. For example, if a person has a state of mood of sadness than this person prefers to listen to a more melancholic song.

The solution does not include complex programming languages; therefore, the proposed solution consumes low resources from current electronic devices. The assessors evaluated an imperceptible interruption regarding the consumed resources in the electronic device.

Perceived parameters of the framework in relation to usability present good results and show the importance of the framework to integrate all these parameters to be a complete tool.

The study presents the sentiment analysis applied to a music recommendation system; however, sentiment metrics could be applied to many other areas.


## ACKNOWLEDGMENT

The authors thank both the Laboratory of Computer Architecture and Networks (LARC) at Escola Politécnica - University of São Paulo and the Computer Science Department at the Federal University of Lavras for the motivation to research in the recommendation system and social network area in multimedia services.


## REFERENCES


[1] R. L. Rosa, D. Z. Rodríguez, and G. Bressan, "SentiMeter-Br: a new social web analysis metric to discover consumers' sentiment," *in Proc. IEEE International Symposium on Consumer Electronics,* Hsinchu, Taiwan, pp. 153-154, Jun. 2013.

[2] R. L. Rosa, D. Z. Rodriguez, and G. Bressan, "Music recommendation system based on user s sentiments extracted from social networks," *in Proc. IEEE International Conference on Consumer Electronics*, Las Vegas, USA, pp. 408-409, Jan. 2015.

[3] Y. Shi, M. A. Larson, and A. Hanjalic, "Towards understanding the challenges facing effective trust-aware recommendation," *in Proc. on Recommender Systems and the Social Web*, Barcelona, Spain, pp. 40-43, Sep. 2010.

[4] J. Healey, R. Picard, and F. Dabek, "A new affect-perceiving interface and its application to personalized music selection," *in Proc. Workshop Perceptual User Interfaces*, Nara, Japan, pp. 4-6, Nov. 1998.

[5] S. Koelstra, C. Muhl, M. Soleymani, J. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: a database for emotion analysis using physiological signals," *IEEE Trans. Affect. Comput*., vol. 3, no. 1, pp. 18-31, Jan. 2012.

[6] K. Yoon, J. Lee, and M.-U. Kim, "Music recommendation system using emotion triggering low-level features," *IEEE Trans. Consumer Electron.,* vol. 58, no. 2, pp. 612-618, May 2012.

[7] C.-M. Chen, M.-F. Tsai, J.-Y. Liu, and Y.-H. Yang, "Using emotional context from article for contextual music recommendation," *in Proc. ACM International Conference on Multimedia*, New York, USA, pp. 649-652, Oct. 2013.

[8] R. Cai, C. Zhang, C. Wang, L. Zhang, and W.-Y. Ma, "MusicSense: contextual music recommendation using emotional allocation modeling," *in Proc. International Conference on Multimedia*, Augsburg, Germany, pp. 553-556, Sep. 2007.

[9] A. G. Crespo, R. C. Palacios, J. M. G. Berbís, and F. G. Sánchez, "SOLAR: social link advanced recommendation system", *Future Generation Computer Systems*, vol. 26, no. 3, pp. 374-380, Mar. 2010.

[10] Y. B. Fernandez, J. P. Arias, A. G. Solla, M. R. Cabrer, and M. L. Nores, "Providing entertainment by content-based filtering and semantic reasoning in intelligent recommender systems," *IEEE Trans. Consumer Electron*., vol. 54, no. 2, pp. 727-735, May 2008.

[11] S. Rendle, "Factorization machines with libFM," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 3, pp. 1-22, May 2012.

[12] F. A□. Nielsen, "A new ANEW: evaluation of a word list for sentiment analysis in microblogs," *in Proc. Workshop on Making Sense of Microposts: Big Things come in Small Packages*, Crete, Greece, pp. 93-98, May 2011.

[13] H. H. Kim, "A semantically enhanced tag-based music recommendation using emotion ontology," *in Proc. Asian Conference on Intelligent Information and Database Systems*, Kuala Lumpur, Malaysia, pp. 119-128, Jan. 2013.

[14] A.C.M. Fong, B. Zhou, S.C. Hui, G.Y. Hong, and T. A. Do, "Web content recommender system based on consumer behavior modeling," *IEEE Trans. Consumer Electron*., vol. 57, no. 2, pp. 962-969, May 2011.

[15] Q. Ye, R. Law, and B. Gu, "The impact of online user reviews on hotel room sales," *International Journal of Hospitality Management*, vol. 28, no. 1, pp. 180-182, Jul. 2009.

[16] R. Feldman, "Techniques and applications for sentiment analysis," *ACM Commun*., vol. 56, no. 4, pp. 82-89, Apr. 2013.

[17] S. Poria, A. Gelbukh, B. Agarwal, E. Cambria, and N. Howard, "Sentic demo: a hybrid concept-level aspect-based sentiment analysis toolkit," *in Proc. European Semantic Web Conference,* pp. 31-35, Crete, Greece, May 2014.

[18] C. M. Whissell, "The Dictionary of Affect in Language," in *Emotion: Theory, Research and Experience. The Measurement of Emotions,* R.Plutchik and H. Kellerman, eds., Academic Press, vol. 4, pp. 113-131, 1989.

[19] G. Paltoglou, and M. Thelwall, "Twitter, MySpace, Digg: unsupervised sentiment analysis in social media," *ACM Trans. on Intelligent Systems and Technology*, vol. 3, no. 4, pp. 1-19, Sep. 2012.

[20] M. Thelwall, D. Wilkinson, and S. Uppal, "Data mining emotion in social network communication: gender differences in MySpace," *J. Am. Soc. Inf. Sci. Technol.,* vol. 61, no. 1, pp. 190-199, Sept. 2010.

[21] D. Yang, and W. Lee, "Music emotion identification from lyrics," *in Proc. IEEE International Symposium on Multimedia*, San Diego, California, USA, pp. 624-629, Dec. 2009.

[22] G. Qiu, F. Zhang, J. Bu, and C. Chen, "Domain specific opinion retrieval," *in Proc. Asia Information Retrieval Symposium on Information Retrieval Technology*, Sapporo, Japan, pp. 318-329, Oct. 2009.

[23] R. L. Rosa, D. Z. Rodriguez, V. A. Souza, and G. Bressan, "Recommendation system based on user profile extracted from an IMS network with emphasis on social network and digital TV," *in Proc. Latin America Networking Conference*, Quito, Ecuador, pp. 40-47, May 2011.

[24] S. E. Shepstone, Z.-H Tan, and S.H Jensen, "Audio-based age and gender identification to enhance the recommendation of TV content," *IEEE Trans. Consumer Electron*., vol. 59, no. 3, pp.721-729, Aug. 2013.

[25] A. C. Lampe, N. Ellison, and C. Steinfield, "A familiar Face(book): profile elements as signals in an online social network," *in Proc. Conference on Human Factors in Computing Systems*, New York, USA, pp. 435-444, May 2007.

[26] A. Yamada, M. Hara, and Y. Miyake, "Exploiting privacy policy conflicts in online social networks," CMU Cylab Technical Report, Feb. 2012.

[27] J. Kunegis, A. Lommatzsch, and C. Bauckhage, "The Slashdot Zoo: mining a social network with negative edges", *in Proc. International Conference on World Wide Web*, New York, USA, pp. 741-750, May 2009.

[28] C. Laurier, M. Sordo, J. Serrà, and P. Herrera, "Music mood representations from social tags," *in Proc. International Conference on Music Information Retrieval*, Kobe, Japan, pp. 381-386, Oct. 2009.

[29] J. A. Russell, "A circumplex model of affect," Journal of Personality and Social Psychology, vol. 39, no. 6, pp. 1161-1178, Dec. 1980.

[30] A. H. Kronbauer, C. A. S. Santos, and V. Vieira, "Smartphone applications usability evaluation: a hybrid model and its implementation," *in Proc. International Conference on Human-Centered Software Engineering*, Toulouse, France, pp. 146-163, Oct. 2012.

[31] X. Gu, Z. Xu, T. Wang, and Y. Fang, "Trusted service application framework on mobile network," *in Proc. International Conference on Ubiquitous Intelligence and Computing*, Fukuoka, Japan, pp. 979-984, Sep. 2012.

[32] D. Z. Rodriguez, R. L. Rosa, E. A. Costa, J. Abrahao, and G. Bressan, "Video quality assessment in video streaming services considering user preference for video content," *IEEE Trans. Consumer Electron*., vol. 60, no. 3, pp. 436-444, Aug. 2014.

[33] Q. Xu, Q. Huang, and Y. Yao, "Online crowdsourcing subjective image quality assessment," *in Proc. ACM International Conference on Multimedia*, Nara, Japan, pp. 359-368, Oct. 2012.

[34] J. M. C. Bastien, and D. Scapin, "Ergonomic criteria for the evaluation of human-computer interfaces," Institut National de Recherche en Informatique et en Automatique, INRIA, Rocquencourt, France, Technical Report 156, 1993.

[35] B. Senach "Ergonomic evaluation of the human-computer interfaces: a review of the literature," Research report, INRIA, Sophia Antipolis, Technical Report 1180, 1990.

[36] P. Rabardel, and P. Béguin, "Instrumented mediated activity: from subject development to anthropocentric design", *Theor. Issues Ergonon. Sci*., vol.7, no. 5, pp. 429-461, Feb. 2005.

[37] R. Likert, "A technique for the measurement of attitudes," *Archives of psychology*, vol. 22, no. 140, pp. 1-55, Jun. 1932.

[38] A. Alexandrov, "Characteristics of single-item measures in Likert scale format," *Electron. Journal of Business Research Methods*, vol. 8, no. 1, pp. 1-12, Sep. 2010.

[39] H. F. Hoffman, and F. Lehner. "Requirements engineering as a success factor in software projects," *IEEE Software*, vol. 18, no. 4, pp. 58-66, Jul. 2001.

## BIOGRAPHIES

**Renata Lopes Rosa** received her M.S. degree from the University of São Paulo - USP (2009). She is a Ph.D. student at Escola Politécnica from the University of São Paulo (EPUSP). Her current research interests include computer networks, quality of experience of multimedia service, social networks and recommendation systems.

**Demóstenes Zegarra Rodríguez** received his M.S. degree (2009) and PhD degree (2013) from the University of São Paulo - USP. He is currently an Adjunct Professor at the Department of Computer Science, Federal University of Lavras, Brazil. He has a solid knowledge in Telecommunication Systems and Computer Science based on 14 years of professional experience in important companies. His research interest includes QoS and QoE in Multimedia services, Digital TV, recommendation systems, social networks and architect solutions in Telecommunication Systems.

**Graça Bressan** got her PhD degree (1986) in Digital Systems at Escola Politécnica from the University of São Paulo (EPUSP). At the Computer Engineering Department of EPUSP, she teaches and develops research on computer network. Her research interests includes computer networks and digital television, distributed systems, QoS mechanisms, middleware for Digital TV, interactive digital TV, recommendation systems, video-conferencing, modeling and performance analysis of networks, and application in distance education.